The Impact of Web-Scale Discovery on the Use of Electronic Resources

Lisa Ngo[a]*, Cody Hennesy[a**] and Ian Knabe[a***]

[a]*UC Berkeley Library, University of California, Berkeley, Berkeley, USA*

*Corresponding author: Engineering Librarian, lngo@berkeley.edu. ORCID: https://orcid.org/0000-0002-7157-6977

**Current Affiliation: Journalism and Digital Media Librarian, University of Minnesota Libraries. ORCID: https://orcid.org/0000-0002-9410-9810

***Current Affiliation: Assistant Head of Research Materials Procurement, University of Houston Libraries. ORCID: https://orcid.org/0000-0003-4715-4878

The Impact of Web-Scale Discovery on the Use of Electronic Resources

In 2015, the University of California, Berkeley, launched EBSCO Discovery Service (EDS), a web-scale discovery tool, with a goal of improving visibility and usage of collections. This study applies linear regression analysis to usage data for ebooks, ejournals, and abstracts and indexing (A&I) databases before and after implementation of EDS in order to identify correlations between the discovery layer and usage of library electronic resources across platforms. Our findings diverge from conclusions drawn in the previous literature that indicate that resource use generally increases after a discovery tool is implemented. We examine data from a longer period of time than the previous literature had, looking for statistically significant changes in resource use. The discovery layer at UC Berkeley did not lead to equal increases across platforms, but rather to a complex array of increases and decreases in use according to a variety of factors.

Keywords: Web-scale discovery; EBSCO Discovery Service; Electronic resources; Academic library, Assessment; Usage data

Introduction

Background

The University of California, Berkeley, is a large research and teaching institution serving over 25,000 undergraduates, 10,000 graduate students, and 1,500 faculty in more than 250 degree programs. The UC Berkeley Library reflects the complexity of the campus with a system that includes more than 20 physically distinct subject branch libraries and several affiliated libraries that are administratively separate from, yet share some resources with the main library system. In addition to a physical collection of more than 12 million volumes, the Library provides access to electronic resources in the form of abstracts and indexing services (A&I databases), ebooks, ejournals, data sets, unique digital collections, as well as the collections of the other nine UC campuses and two storage facilities. User access to the physical and electronic collections require

navigation across disparate information silos including multiple OPACs, vendor-developed interfaces, and home-grown systems.

In 2012 a Web Services Review Team was charged with reviewing the existing library website to make recommendations for a redesign that would produce a more user-friendly and intuitive site. During user testing conducted by the review team between 2012-2014 it became increasingly clear that navigating multiple discrete information silos presented a major barrier to access for many students attempting to locate information resources. Even after a new library website was launched in 2014, these barriers, though slightly lowered due to a more streamlined design, continued to persist. It was determined that a web-scale discovery layer was a solution that could address this issue within a reasonable development timeline and without overtaxing staff workloads. Most importantly, a discovery layer would provide a single-search-box experience, which local user testing during the website redesign process showed students overwhelmingly expected and preferred.

Web-scale discovery services have been available since the early 2000s, with multiple commercial products entering the market alongside a number of open source projects developed by libraries (Breeding, 2010). UC Berkeley had experience with metasearch tools, having deployed the California Digital Library's federated search tool SearchLight in the 2000s. SearchLight was retired in 2005 after it was determined that the system was not robust enough to meet the needs of the the UC system (California Digital Library, 2005). Indeed, federated search tools notoriously suffered from slow response times and time-outs (Thomsett-Scott & Reese, 2012), problems that a new generation of web-scale discovery services have largely overcome (Vaughan, 2011). At the time of UC Berkeley's evaluation, four services dominated the discovery market: Ex Libris Primo, EBSCO Discovery Services (EDS), ProQuest Summon, and OCLC

WorldCat Discovery Service. After an evaluation of all four products, EDS was selected and launched on the UC Berkeley campus in Fall 2015. Improving visibility and access to library collections (both physical and electronic) was a core goal, second only to improving usability of library search.

Literature Review

In the process of investigating discovery services, the EDS implementation team was particularly interested in the potential impact on the use of licensed electronic resources across platforms. It was evident from previous literature that citation data available in a central index can expose users to more licensed online content (Vaughan, 2011). We also know that rich bibliographic metadata increase the findability and use of physical library resources (Tosaka, Weng, & Weng, 2011), and the discovery tool would allow us to return library catalog records alongside results from the central index. In one of the earliest studies of resource usage statistics before and after the implementation of a discovery service, Doug Way (2010) examined COUNTER statistics for database searches and full-text downloads for the year before and after implementation of Serials Solutions' (later, Proquest) Summon product. He found a dramatic increase in full-text downloads from library subscriptions immediately after implementation at Grand Valley State University, along with a decrease in use of abstract and indexing databases, noting that the trends stood across disciplines and platforms.

Tony Greiner (2011) also found an increase in full text downloads at libraries in the Orbis Cascade Alliance after their WorldCat Local implementation. Lisa O'Hara (2012) reported similar results after the implementation of Summon at The University of Manitoba Libraries, as did Jan Kemp (2012) at the University of Texas at San Antonio (using Summon), and Kristin Calvert (2015) at Western Carolina University (using EDS). Calvert, O'Hara, and Way all used COUNTER statistics in their analysis

"to ensure that measurements for electronic resources were comparable" (O'Hara, 2012, p.27). Kemp looked at full-text click-throughs from the link resolver as a measure of full text use, and Greiner stated "only those databases that provide reliable use data are included here" (Greiner, 2011, p.214), with no further information about whether COUNTER statistics were specifically used.

In contrast, Zebulin Evelhoch (2016) evaluated COUNTER statistics before and after the implementation of Primo at Central Washington University and found a decrease in database searches as well as full-text views. Evelhoch posits that students might have been finding more freely available resources online via Google, skipping the library altogether, or "that users may be unclear how to view or request journal articles using the Primo interface and abandoned their search altogether" (Evelhoch, 2016, p.206). Evelhoch could not definitively say whether Primo had any effect on usage, but suggested further research.

Each of these studies compared at least one year of pre-implementation data with at least one year of post-implementation data. O'Hara (2012) collected data from one year prior to implementation and two years after, while Calvert (2015) looked at three semesters pre-implementation and three semesters post-implementation, and Evelhoch (2016) compared two years of pre-implementation data with one year of post-implementation data.

While there is a great deal of literature on the usability and relevance of search results in EDS[1], only two previous studies have looked for correlations between the implementation of EDS specifically and electronic resource use. Calvert (2015) finds that ejournal use overall increased and circulation of print materials decreased following an implementation of EDS at Western Carolina University, though without testing the

[1] See, for example, Williams and Foster, 2011; Asher et al., 2013; Bonner and Williams, 2016.

statistical significance of the changes detected. Angela Pratesi (2018) examined COUNTER reports for two databases—RILM Abstracts of Music Literature and Music Periodicals Database—as her institution implemented EDS but found that changes to COUNTER reports and definitions during the time period under study significantly obfuscated any longitudinal usage trends.

In addition to individual case studies, Michael Levine-Clark, Jason Price, and John McDonald (2014) and Valerie Spezi, Claire Creaser, and Angela Conyers (2015) studied usage data gathered across academic institutions. The Levine-Clark et al. study is particularly interesting as it studies larger-scale usage changes across libraries, including changes across libraries using the same discovery service, and the effect of discovery service across publishers. They found that usage of specific publisher content was affected—some negatively and some positively—by the discovery service used, and that libraries with a discovery service in place generally had a greater increase in usage of publisher hosted journal content than libraries without one. In another study of discovery services at multiple institutions, Eugene Barsky, Sarah Jane Dooley, Tara Mawhinney, Zoey Peterson, and Michelle Spence (2013) focused on science and engineering ebook usage at four Canadian institutions, two with Summon implementations and two with WorldCat implementations, and found that usage varied significantly depending on local implementation of the discovery service and level of indexing for ebooks.

Our study uses data from a single institution, UC Berkeley, and we contribute new insights about the interactions between discovery services and library usage along the following dimensions. We take a longer view of usage trends before and after discovery implementation than is found in the existing literature, evaluating two years of data from before the EDS implementation with two years of data post-

implementation. We take a granular view of electronic resources by comparing use for three different categories: ejournals, ebooks, and abstract and indexing databases (both those included in EDS and not included in EDS). Finally, instead of analyzing usage trends based on percentage change calculations in usage statistics, we apply linear regression analyses to the data to test for statistical significance in usage changes before and after discovery implementation.

Method

To look for correlations between changes in electronic resource usage and discovery implementation, we collected 48 months of data from publishers for ejournals and ebooks and 30 months of data from abstract & indexing databases. We only include vendors that were able to provide COUNTER compliant data in order to be consistent in how use was measured (see Appendix for more information about COUNTER data). The 48 months of data was collected for four academic years, spanning from July 2013 to June 2017; for A&I databases, 30 months of data spanned from January 2015 to June 2017. EDS was launched in July 2015, producing two years of data prior to launch and two years of data post-launch for ejournals and ebooks and six months of data prior to launch and two years of data post-launch for A&I databases. Our study does not take into account changes in content availability within platforms or publishers (e.g., large increases or decreases in number of titles and/or content) due to unavailability of data from most publishers and inaccuracies in data that was made available.

Ejournals

Usage data for publishers of ejournals was gathered from JR1 reports, using counts of successful full-text article requests regardless of file format. We only examined JR1 reports, and did not differentiate between JR1 and JR1a or JR1GOA, as Open Access or non-current use was immaterial to our investigation.

Ebooks

Usage data for publishers of ebooks was gathered from BR2 reports, using counts of successful section requests. BR2 reports were utilized since the majority of publishers make ebook content available for download at the "section" level. These sections can range from book chapters to dictionary entries, and vary widely in length; despite these variations, BR2 section requests provided a reasonable metric to track longitudinal use. BR1 reports were not used since those reports had not been regularly collected at Berkeley, leading to a lack of historical reports available for analysis, and as previously mentioned, the majority of our ebooks and ebook platforms report use at the section level.

Abstract & Indexing databases

Usage data for A&I databases was gathered from DB1 reports measuring Result Clicks. Result clicks were introduced as a metric with COUNTER release 4 for measuring usage in A&I databases. However, we note that the Result Clicks metric was not standardized in COUNTER statistics until 2014, and not all databases we examined provided this data consistently until January 2015. As a result, our data for A&I databases span from January 2015 to June 2017 and does not include data from July 2013 to December 2014. While this shortened our sample range, we felt that usage clicks represented a clearer metric for assessment of use, as it captures user-selected responsive search results independent of platform, while searches might be artificially elevated by users' query testing and refinements.

Statistical Analysis

After the data was collected, outliers for each publisher or database were identified and removed using the standard formula of calculating the Interquartile Range (IQR) of the variables and removing values that were 1.5IQR greater than the upper quantile and

1.5IQR less than the lower quantile. Outliers are assumed to be present in the reports due to incidents of systematic downloading during breaches or other incidents involving compromised user credentials.

Multiple linear regression analysis was used to determine whether the implementation of our discovery service correlated with resource usage. To account for the natural peaks and valleys of resource use during the regular semester, we used semester "peaks" as an independent variable in our calculations in addition to discovery as another independent variable.

The multiple linear regression analysis equation used for our study is as follows:

$$y_t = \beta_0 + \beta_1 P_1 + \beta_1 P_2$$

The dependent variable, y, equals the number of article downloads (for ejournals), or number of chapter downloads (for ebooks), or number of result clicks (for A&I databases).

Our independent variables are P_1 and P_2. P_1 indicates whether usage was before Discovery launch or after Discovery launch. We set $P_1 = 0$ before Discovery launch, and $P_1 = 1$ post-Discovery launch. Our second independent variable is P_2, indicating whether usage was during semester "peaks" or not during "peaks". Semester "peaks" are defined as the high use months of January, February, March, April, May, September, October, and November. $P_2 = 0$ for months outside of the "peaks" period, and $P_2 = 1$ for months during semester "peaks". Our null hypothesis assumes that the launch of a discovery service has no effect on usage of the library's resources.

Data was analyzed using IBM SPSS Statistics software. Our raw data can be accessed at https://doi.org/10.6078/D13D4V.

Results

Ejournals

Data was gathered and analyzed for sixteen ejournal platforms, including single and multi-publisher platforms, society, and university presses. Figure 1 shows the usage for analyzed platforms where usage changes correlated (positively or negatively) with EDS, with outliers removed and null values plotted. EDS launched in July 2015 and is indicated by the black line. Figure 2 shows the usage for analyzed platforms with usage changes that showed no correlation with EDS.

(insert Figure 1)

(insert Figure 2)

Table 1 shows the results of the linear regression analysis and lists p-values and coefficients for platforms with use effected by EDS. Platforms with usage changes negatively correlated with EDS are highlighted. Table 2 shows the p-values and coefficients for platforms where usage changes showed no correlation with EDS.

(insert Table 1)

(insert Table 2)

Over half of the ejournal platforms— eleven out of sixteen—had p-values that led to rejection of the null hypothesis (p-value < 0.05), meaning that the implementation of EDS was associated with changes in resource usage for these platforms (see Table 1). For the remaining five publishers (see Table 2), there was no observable correlation between the implementation of EDS and resource usage (p-value > 0.05).

Of the eleven platforms where a correlation between EDS and resource usage was evident, five had positive coefficients, showing that usage increased after EDS was implemented. These tended to be larger publishers—Taylor & Francis, Wiley, Elsevier, and IEEE—though Gale is a smaller platform that showed increased use. Gale and

Taylor & Francis showed high coefficients, which illustrates that usage increased significantly on these two platforms following the implementation of EDS. Six publishers, however, had corresponding negative coefficients, showing that usage decreased after EDS was implemented. These resources included a mix of large, medium, and small publishers, with JSTOR and Highwire being the largest. Society publishers such as Cambridge and ACM are also included in this category.

EBooks

Data was gathered and analyzed for nine ebook platforms, including single and multi-publisher platforms, society, and university presses. The analysis includes data from Ebrary (now Proquest Ebooks Central), links to which are *not* included directly in EDS search results but are represented in library catalog records that appear in EDS.

Figure 3 shows usage for ebook platforms with outliers removed and null values plotted. The EDS launch date is indicated by the black line.

(insert Figure 3)

Five of the nine ebook platforms analyzed had p-values that showed correlation between resource usage and EDS implementation, and four of those five had positive coefficients, showing that usage increased after the implementation of EDS (see Table 3). One publisher, Oxford University Press, had a negative coefficient, showing that use dropped after EDS was implemented. The remaining four publishers (represented in Table 4) showed no correlation between EDS implementation and usage (p-value > 0.05).

(insert Table 3)

(insert Table 4)

Abstracts and Indexing Databases

Data was gathered and analyzed for twelve A&I databases, including databases loaded as part of the EDS index and databases not indexed in EDS, for January 2015 to June 2017. Figure 4 shows the usage for A&I databases analyzed, as measured by result clicks, with outliers removed and null values plotted. Databases not indexed by EDS are indicated with "(ND)". The EDS launch date is indicated by the black line.

(insert Figure 4)

Table 5 shows the results of the linear regression analysis and lists p-values and coefficients for A&I databases with use changes correlated with EDS. Table 6 shows the p-values and coefficients for A&I databases where usage is not correlated with EDS.

(insert Table 5)

(insert Table 6)

Only two out of the twelve A&I databases analyzed showed statistically significant change in usage following EDS implementation: Georef, which is not included in the EDS index, and MathSciNet, which is included in the EDS index. Georef had a p-value < 0.05 but a negative coefficient, showing that use decreased after EDS was implemented. MathSciNet also had a p-value < 0.05 but had a positive coefficient, showing an increase in use following the implementation of EDS. The remaining ten A&I databases showed no statistically significant change in usage, regardless of whether their content was indexed in EDS.

Discussion

Ejournals

While the implementation of EDS at UC Berkeley positively correlates with increased use of ejournals on the majority of the platforms we examined, it did not appear to be positively correlated with more successful full-text article requests (as measured by JR1

reports) across platforms equally. Rather, it was associated with decreases in full-text article requests on more platforms (six) than those for which it showed an increase in requests (five). These results differed from both the Levine-Clark et. al. (2014) and Calvert (2015) studies that found increased usage for a majority of publishers at institutions with EDS. In fact, of the ten platforms examined by both Calvert and this study, only three shared similar usage trends (positive correlation between EDS and downloads on the Elsevier platform; no correlation between EDS and downloads on the Springer platform; and negative correlation between EDS and downloads on the Project Muse platform). The remaining seven platforms showed conflicting usage trends between the two studies. Given this variance in results, it's unfortunate that Levine-Clark chose not to name the publishers tracked in their study; of the six publishers they examined, four showed increased usage with EDS while two showed decreased usage, and though we don't know who the publishers are it is interesting to note that one publisher in particular showed decreased usage consistently across three different discovery products.

Our investigation found that records from four of the six platforms showing decreased use at UC Berkeley are included in EDS indexes but are not available to include as "a database" in EDS. When a library enables access to a subscribed full text platform as a database in EDS, direct links to the full-text on the platform appear in the EDS search results. Citations from platforms that are in the central index but not available as a database in EDS, however, are more likely to appear in EDS search results with a link to UC Berkeley's OpenURL link resolver, UC-eLinks[2]. Mary Ann

[2] The precise level of inclusion of content in the EDS indexes varies significantly based on each platform or publisher's inclusion in EBSCO's "Complementary," "Supplemental," or other form of index, as well as the institution's subscription status to a particular resource. See EBSCO Help for more

Liebert, ACM, Cambridge, and Highwire[3] all appear in the EDS indexes, but citations from those resources in the EDS search results do not link directly to full-text on their respective platforms. Users who select the UC-eLinks option from an EDS search result from these platforms might find their way to the platform after a few clicks, but the implementation of EDS at UC Berkeley appears to be negatively correlated with use of those resources. Based on this data, it is clear that inclusion in the EDS indexes does not necessarily correlate with an increase in full-text article requests.

Two other major platforms showing fewer full-text article requests after the implementation of EDS, Project Muse and JSTOR, are both brands with substantial name recognition among our primary audiences at Berkeley. That name recognition might have contributed to comparatively high use before EDS implementation, which then adjusted downward as EDS became available and directed users to a more diverse array of electronic resources. We have observed informally that JSTOR, for example, is one of a few scholarly article platforms that students in the humanities and social sciences recognize by name when they arrive on campus. Those students, prior to the EDS implementation, might have been more likely to go straight to JSTOR when they needed journal articles for their research. Following implementation, the EDS search box, which is highlighted prominently on the library website homepage[4], provided an alternative user-friendly source for novice users looking for journal articles. If some

---

information: https://help.ebsco.com/interfaces/EBSCO_Discovery_Service/EDS_Admin_Guide/content_included_in_EDS_profile

[3] The steep decrease in use of Highwire was likely also due to a platform shift effective December 2016, when Sage ejournals moved to their own platform. While it was not feasible to track publication and publisher changes across platforms, it may be the case that major changes such as the shift from Highwire to Sage complicate the story that our regression analysis tells.

[4] To view the UC Berkeley Library website as of September 21, 2018, see: https://web.archive.org/web/20180921070037/http://lib.berkeley.edu

subset of students now use EDS instead of going directly to JSTOR, it stands to reason that JSTOR use would fall, even as JSTOR articles are available in EDS. It is also worth noting that at the same time EDS was launched the library website was substantially redesigned, shifting users' paths to specific electronic resources such as JSTOR. While users retained the ability to search native interfaces directly, their access points changed parallel to the appearance of the EDS search box. Overall, however, the usage shifts seen for JSTOR and Project Muse suggest that users are discovering relevant sources from a more diverse range of platforms following the implementation of EDS.

Conversely, Elsevier, IEEE, Wiley, Gale, and Taylor & Francis all saw full-text article requests rise after the implementation of EDS. Notably, significant subsets of resources from two of these platforms, Elsevier and IEEE, are included in EDS "as databases." ScienceDirect (Elsevier), for example, is enabled as a database in the local implementation of EDS. Citations from these platforms are either available in the EDS search results as full-text links to content on the originating platform or as full-text within the EDS interface itself. This form of inclusion can also mean that resources from these platforms are available for full-text searching via EDS. It's likely that the close integration of resources from Elsevier and IEEE into EDS search results positively affected use.

Ejournal packages available from Wiley, Gale, and Taylor & Francis are either not available to include or have not been enabled "as databases" in EDS at Berkeley, but still these platforms saw meaningful increases in use. In fact, Gale and Taylor & Francis saw the largest increases of all of the ejournal platforms under scrutiny. This suggests, first, that ejournal platforms that are included in the EDS indexes without integration "as databases" can still see increased use. Second, all five platforms in this category do not generally have strong brand-recognition as go-to interfaces for scholarly content,

compared especially to a platform like JSTOR. These platforms may be more often known as publishers (IEEE), as conglomerates of publisher content (Taylor & Francis), or as conglomerates of databases and other access points (Elsevier), if they are recognized at all by users. Unlike JSTOR, then, these platforms might have had more to gain from the implementation of an interface that reduces the need for a student to actively choose a branded entry-point for research.

Of the five platforms that were not affected by the EDS implementation—Ingenta, Oxford, ACS, Ovid, and Springer—none are strictly included or enabled in UC Berkeley's EDS implementation "as databases." There may then also be a middle-ground for platform content that is not available in EDS as direct full-text or full-text links.

Ebooks

EDS implementation positively correlated with higher rates of successful ebook section requests, as tracked in BR2 reports, on four platforms, and negatively with use on one platform (Oxford). Two of the platforms where ebook usage increased (Elsevier and Wiley) also showed increased use for ejournals, while another platform (Cambridge) showed increased ebook access but decreased access of ejournal content. The fact that different formats were impacted differently on Cambridge suggests that format differences are meaningful, and that usage shifts arising after EDS implementation are not entirely dependent on platform.

It is possible that ebook use increased more consistently than ejournal use due to a kind of double-availability of ebooks in EDS: UC Berkeley library catalog (OskiCat) records for ebooks appear in EDS search results, and many ebook results also appear in EDS as records from specific platforms or as "Online access" links derived from MARC 856 fields in the catalog record. Most often, duplicate records are merged in the EDS

search results, offering an online direct full-text link as well as a link to the OskiCat record under the heading of a specific book title. This could explain why the use of Ebrary (now ProQuest Ebooks Central) ebooks showed no statistically significant change, despite the fact that Ebrary ebook records in EDS search results do not include direct full-text links to ProQuest records. Users are still able to find direct links to the full-text of these ebooks in EDS via OskiCat records. This supports the conclusion of Barskey et. al. (2013) that the inclusion of ebooks in a discovery layer, whether full-text indexed or not, can positively impact usage.

Ebooks from EBSCO, the discovery platform's parent company, saw the largest usage increase of all ebook platforms observed in the study. This is unsurprising, particularly considering that Berkeley has configured EDS search results to prioritize display of the most reliable full-text content (and verified full-text links) available for any given title. EBSCO is well-equipped to verify the availability of full-text ebooks from their own platform, and in cases where ebooks are available at Berkeley from more than one platform (Springer and EBSCO, for example) only the most reliable full-text link will display (EBSCO, rather than Springer, in this example). ScienceDirect (Elsevier), also stands out as an ebook platform where use increased, and is one of the only ebook platforms included, not only in the EDS indexes, but also "as a database" in EDS.

Springer stands out in the analysis as a large publisher that did not see statistically significant changes in either ebook or ejournal use following the implementation of EDS, in contrast to usage at the University of British Columbia and at the University of Toronto where Barsky et. al. (2013) reported dramatic increases after Summon implementation, and at the University of Manitoba, where O'Hara (2012) saw a greater than 100% increase in Springer ebook usage post-Summon. Neither

Springer ebooks or ejournals are included "as a database" in EDS, but subscribed Springer content is available in the EDS complementary index, and Springer ebooks often surface as full-text links from the library catalog. Still, there was no observable change in use of ejournal or ebook content from the Springer platform. Nor were detectable changes observed for ebooks from ACM or CABI. Oxford University Press was the only ebook platform that saw decreased access following the implementation of EDS. Springer was able to provide a non-COUNTER report with information on referring URLs from Berkeley users accessing SpringerNature content indicating a majority of users navigated to SpringerNature content from Google or Google Scholar. Due to lacking availability of similar reports from other publishers/vendors, we were not able to more thoroughly investigate this angle.

A & I databases

Contrary to findings in the literature, we did not see a statistically significant difference in the usage of A&I databases after the discovery layer was launched. Only one of the twelve databases analyzed showed increased usage after EDS implementation, one database showed decreased usage after EDS implementation, and the remaining ten databases showed no change at all, regardless of whether the content was indexed in EDS; in contrast, Calvert (2015) found usage increases in the majority of A&I databases indexed in EDS. This was surprising since we expected that students would stop searching individual subject databases and start using EDS as their primary search interface, thus decreasing individual database usage. In particular we expected usage of Compendex and PsycInfo, two databases not included in EDS indexes, to decrease following the EDS implementation.

The discrepancy between our findings and those of other authors might stem in part from our choice to use result clicks as a metric rather than the more commonly used

"searches." We note that Evelhoch (2016) evaluated result clicks and record views and found similar effects on database usage. Given that some librarians at UC Berkeley found the result click data to be surprisingly low, we question whether that data was accurately measured in the DB1 reports and consequently whether it was the best metric for our purposes.

If we accept that the result clicks data are accurate and reflective of actual use, however, we can conclude that there was no correlation between EDS and usage of A&I databases at UC Berkeley. The ways in which discovery is taught (or often *not* taught) by subject liaisons and reference staff at Berkeley might explain this in part (for this particular study, we did not take into account whether faculty, lecturers, and other teaching staff changed the resources to which they directed students). Many librarians chose not to teach discovery when visiting classes or at the reference desk, but instead continued to direct students to subject-specific platforms for their research. Librarians for engineering, life and health sciences, psychology, history and art history—all heavy users of Compendex, PsycInfo, MLA, Avery Index to Architectural Periodicals, Historical Abstracts, INSPEC, BIOSIS, and/or Web of Science—continued teaching A&I databases in their course-integrated instruction, possibly stabilizing usage post-EDS implementation.

Future research might evaluate whether searches, sessions, result clicks, or other measurements could more accurately reflect usage of A&I databases.

Conclusion and Future Directions

Given that previous literature demonstrates a consistent trend of dramatic increases in resource use after the implementation of a discovery layer, it was interesting to note that the impact at UC Berkeley was less clear. While the use of some platforms and formats clearly (and sometimes dramatically) increased following the launch of EDS, the use of

resources on other platforms and formats clearly decreased, and others remained static. The discovery layer at UC Berkeley did not lead to equal increases across platforms, but rather to a complex array of increases and decreases in use according to a variety of factors. Our method of analysis, including longer periods of usage and looking for statistically significant changes in the data, resulted in conclusions that differed from the findings in the current literature and argues for deeper analysis of usage data.

Adding to the complexity of the picture, several issues arose during our study that may require further discussion in future studies of this kind.

Quality of records

The scope of our analysis did not allow for a close inspection of the impact that specific metadata elements in an EDS record have on that record's appearance in the EDS search results, but our results suggest that the level and quality of indexing does effect usage on the ejournal and ebook level. Given that the literature also indicates that the quality of metadata impacts resource usage (Tosaka et al., 2011), institutions wishing to increase use may consider a closer examination of metadata richness and its effect on findability and relevancy ranking in a discovery system.

Availability of full-text in EBSCO

For ejournals, it is difficult if not impossible to tell whether the inclusion of full text directly in the EDS interface effects download statistics from publisher platforms. Our analysis of both ejournals and ebooks suggests that full text availability in the EDS interface may in fact negatively correlate with usage reported by a publisher or platform. In particular, further analysis could be done examining whether the content in ejournal platforms that showed a statistically significant decrease in use after discovery launch—Highwire, Mary Ann Liebert, Project Muse, ACM, JSTOR, and Cambridge—is duplicated in EBSCOhost and available for full text download within the EDS

interface. The concern around biased results or preferential treatment of certain providers in discovery services is not new, nor is anecdotal suspicion that discovery systems created by content providers may be preferentially delivering their own content (Breeding, 2015). A possible avenue for further investigation may utilize COUNTER JR5 reports, which count usage based on year of article publication. Taking into account any embargo periods, usage by year of publication data may illustrate users being directed away from the publisher platform for specific years in which EBSCOhost holds full text content, or in looking for drops in usage on the publisher platform (and correlated spikes on the EBSCOhost platform) as content falls out of the publisher embargo period and becomes available on EBSCOhost.

Access Trends

The usage statistics analyzed in this study were supplied by publisher reports of downloads from their own platforms. It's unclear from these reports exactly how users arrived at a particular resource, whether through a link resolver, Google Scholar, direct links, or other avenues. While some publishers readily provide this information in the form of source reports or referring URL reports, many others do not make this information available. More available data in this area would allow researchers to study shifting trends in how users find ebooks and articles, and to more clearly examine how usage changes on a particular platform may be attributed to a discovery tool or other factors beyond the library's interface. Additionally, trends in syndication and distributed usage logging also stand to heavily impact access and usage tracking beyond publisher platforms.

Circulation of physical collections

Unfortunately, the Integrated Library System used to manage check-outs from the physical collections at UC Berkeley discards usage statistics older than three years, and

we did not consistently collect data early enough, or for a long enough period of time, to tell an accurate story of check-outs before and after EDS implementation. Comparing ebook increases to physical check-outs would have allowed us to consider the relationship between the discovery layer and the use of the collection more holistically. Access to longer term check-out data would also have made it possible to consider the impact of EDS on the use of our physical collections.

Acknowledgements

The authors would like to thank the following individuals and groups at UC Berkeley who helped us to collect and decipher usage statistics in a complex environment: Jason Dezember, Lynne Grigsby, and Sherry Lochhaas from the Library; Lei Zhang, Fan Dong, and the Department of Statistics Statistical Consulting Service.

References

Asher, A., Duke, L.M., & Wilson, S. (2013). Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources. *College & Research Libraries, 74*(5), 464-488. https://doi.org/10.5860/crl-374

Barsky, E., Dooley, S. J., Mawhinney, T., Peterson, Z., & Spence, Z. (2013, June). Influence of Discovery Search Tools on Science and Engineering e-books Usage. Session W435. *The 120th Annual American Society for Engineering Education Conference, June 26, 2013*. Atlanta, GA: American Society for Engineering Education. https://doi.org/10.14288/1.0077942

Bonner, S., & Williams, G. (2016). A Small Academic Library and the Power of EBSCO Discovery Service. *Serials Review*, 42(3), 187-191. https://doi.org/10.1080/00987913.2016.1205428

Breeding, M. (2010). The state of the art in library discovery 2010. *Computers in Libraries, 30*(1), 31–34. https://doi.org/Article

Breeding, M. (2015). Discovery product functionality. *Library technology reports, 50*(1), 5-32. https://doi.org/10.5860/ltr.50n1

California Digital Library. (2005). SearchLight: Lights Out! Retrieved October 17, 2018, from https://www.cdlib.org/cdlinfo/2005/08/25/searchlight-lights-out/

Calvert, K. (2015). Maximizing Academic Library Collections: Measuring Changes in Use Patterns Owing to EBSCO Discovery Service. *College & Research Libraries, 76*(1), 81–99. https://doi.org/10.5860/crl.76.1.81

Evelhoch, Z. (2016). Web-Scale Discovery: Impact on Library Database Web Page Views and Usage. *Journal of Web Librarianship*, *10*(3), 197–209. https://doi.org/10.1080/19322909.2016.1191048

Greiner, T. (2011). How Does Switching to a Discovery Tool Affect Circulation? In *Declaration of Interdependence: The Proceedings of the ACRL 2011 Conference, March 30− April 2* (pp. 211–218).

Kemp, J. (2012). Does Web-Scale Discovery Make a Difference?: Changes in Collections Use after Implementing Summon. In *Planning and Implementing Resource Discovery Tools in Academic Libraries* (pp. 456–468). https://doi.org/10.4018/978-1-4666-1821-3.ch026

Levine-Clark, M., Mcdonald, J., & Price, J. S. (2014). The effect of discovery systems on online journal usage: a longitudinal study. *Insights: The UKSG Journal*, *27*(3), 249–256. https://doi.org/10.1629/2048-7754.153

O'Hara, L. (2012). Collection usage pre- and post-summon implementation at the University of Manitoba Libraries. *Evidence Based Library and Information Practice*, *7*(4), 25–34. https://doi.org/10.18438/B8DK70

Pratesi, A. (2018). Use(less) Data: Discovery, COUNTER, and Music Databases. *Music Reference Services Quarterly*, 21(4), 171-184. https://doi.org/10.1080/10588167.2018.1527276

Spezi, V., Creaser, C., & Conyers, A. (2015). The impact of RDS on usage of electronic content in UK academic libraries: Selected results from a UKSG-funded project. *Serials Review*, *41*(2), 85–99. https://doi.org/10.1080/00987913.2015.1035991

Thomsett-Scott, B., & Reese, P. E. (2012). Academic Libraries and Discovery Tools: A Survey of the Literature. *College & Undergraduate Libraries*, *19*(2–4), 123–143. https://doi.org/10.1080/10691316.2012.697009

Tosaka, Y., Weng, C., & Weng, C. (2011). Reexamining Content-Enriched Access: Its Effect on Usage and Discovery. *College & Research Libraries*, *72*(5), 412–427. https://doi.org/10.5860/crl-137

Vaughan, J. (2011). Web Scale Discovery Services. *Library Technology Reports*, *47*(1), 5–12. https://doi.org/10.4018/978-1-4666-1821-3.ch038

Way, D. (2010). The Impact of Web-scale Discovery on the Use of a Library Collection. *Serials Review*, *36*(4), 214–220. https://doi.org/10.1016/j.serrev.2010.07.002

Williams, S., & Foster, A. (2011). Promise Fulfilled? An EBSCO Discovery Service

Usability Study. *Journal of Web Librarianship*, 5(3), 179-198. https://doi.org/10.1080/19322909.2011.597590

Appendix

COUNTER and COUNTER Report types, explained further

COUNTER is a non-profit organization that provides a Code of Practice for consistency in the reporting of electronic resource usage, seeking to standardize usage reports across vendors. This standardization allows libraries to more accurately compare use across vendors or platforms that are COUNTER-compliant. For the reports analyzed in this paper, Release 4 of the Code of Practice was the most current standard. COUNTER not only provides for consistent metrics, but additionally a standard format for the reports.

JR1 report: The COUNTER JR1 report provides a monthly count of full-text article requests by title. The JR1 Report includes any articles published under that title, while JR1a and JR1 GOA provide more specific counts: the JR1a only counts archival or backfile content, and the JR1 GOA only counts downloads for Gold Open Access articles.

BR2 Report: The COUNTER BR2 report provides a monthly count of section requests by title, with section requests generally meaning book chapters. In contrast, the BR1 COUNTER report counts full book requests.

DB1 Report: The COUNTER DB1 (Release 4) report provides monthly counts across three metrics: Searches, Result Clicks, and Record Views. As mentioned previously, Release 4 added Result Clicks and Record Views in 2014, with consistent adoption across the reports analyzed in January 2015, shortening the period for which we were able to analyze compared to Journals and Books. Result Clicks counts each click from a set of search results, which could be viewing a detailed record, or viewing full text (either on that platform, or linking out to full text on a different platform), while Record Views counts when users view an abstract, detailed metadata, or the full database record.

DB1 reports, and how Result Clicks and Record Views are counted (and how those relate to full text requests) can be quite confusing; Athena Hoeppner's "Explaining COUNTER r4" slide (https://www.slideshare.net/AthenaHoeppner/explaining-counter4) shows how these counts can accumulate across platforms and report types.

Table 1. Linear regression analysis of ejournal platforms showing statistically significant changes following EDS implementation (p-value < 0.05)

Ejournal Platform	**p-value**	**Coefficient**
Highwire	0.000	-0.783
Mary Ann Liebert	0.001	-0.461
Project Muse	0.005	-0.379
Association of Computing Machinery (ACM)	0.022	-0.289
Cambridge	0.042	-0.248
JSTOR	0.050	-0.269
Wiley	0.001	0.414
Elsevier	0.001	0.430
Institute of Electrical and Electronics Engineers (IEEE)	0.000	0.524
Gale	0.000	0.724
Taylor & Francis	0.000	0.780

Table 2. Linear regression analysis of ejournal platforms that did not show statistically significant changes following EDS implementation (p-value > 0.05)

Ejournal Platform	**p-value**	**Coefficient**
American Chemical Society (ACS)	0.227	0.171
Ingenta	0.287	-0.168
Oxford University Press (OUP)	0.383	0.110
Ovid	0.157	0.224
Springer	0.081	0.280

Table 3. Linear regression analysis of ebook platforms showing statistically significant changes following EDS implementation (p-value < 0.05)

Ebook Platform	**p-value**	**Coefficient**
Oxford University Press	0.002	-0.370
Wiley	0.049	0.288
ScienceDirect (Elsevier)	0.002	0.435
Cambridge	0.000	0.606
EBSCO	0.000	0.707

Table 4. Linear regression analysis of ebook platforms that did not show statistically significant changes following EDS implementation (p-value > 0.05)

Ebook Platform	**p-value**	**Coefficient**
Association of Computing Machinery	0.143	-0.233
Centre for Agriculture and Bioscience International (CABI)	0.931	0.013
Ebrary (ProQuest Ebooks Central)	0.971	-0.004
Springer	0.463	0.114

Table 5. Linear regression analysis of A&I databases showing statistically significant changes following EDS implementation (p-value < 0.05)

A&I Database	**Indexed in EDS**	**p-value**	**Coefficient**
Georef	No	0.038	-0.392
MathSciNet	Yes	0.011	0.453

Table 6. Linear regression analysis of A&I databases showing no statistically significant changes following EDS implementation (p-value > 0.05)

A&I Database	**Indexed in EDS**	**p-value**	**Coefficient**
Academic Search Complete	Yes	0.35	0.167
Art Index Retrospective	Yes	0.256	0.186
Avery Index to Architectural Periodicals	Yes	0.1	-0.26
BIOSIS	Yes	0.985	0.004
Compendex	No	0.068	0.323
Historical Abstracts	Yes	0.16	0.236
INSPEC	Yes	0.225	0.227
MLA	Yes	0.406	-0.159
PsycInfo	No	0.642	0.098
Web of Science	Yes	0.292	0.2

Figure 1. Ejournal platforms where usage changes correlated (positively or negatively) with EDS

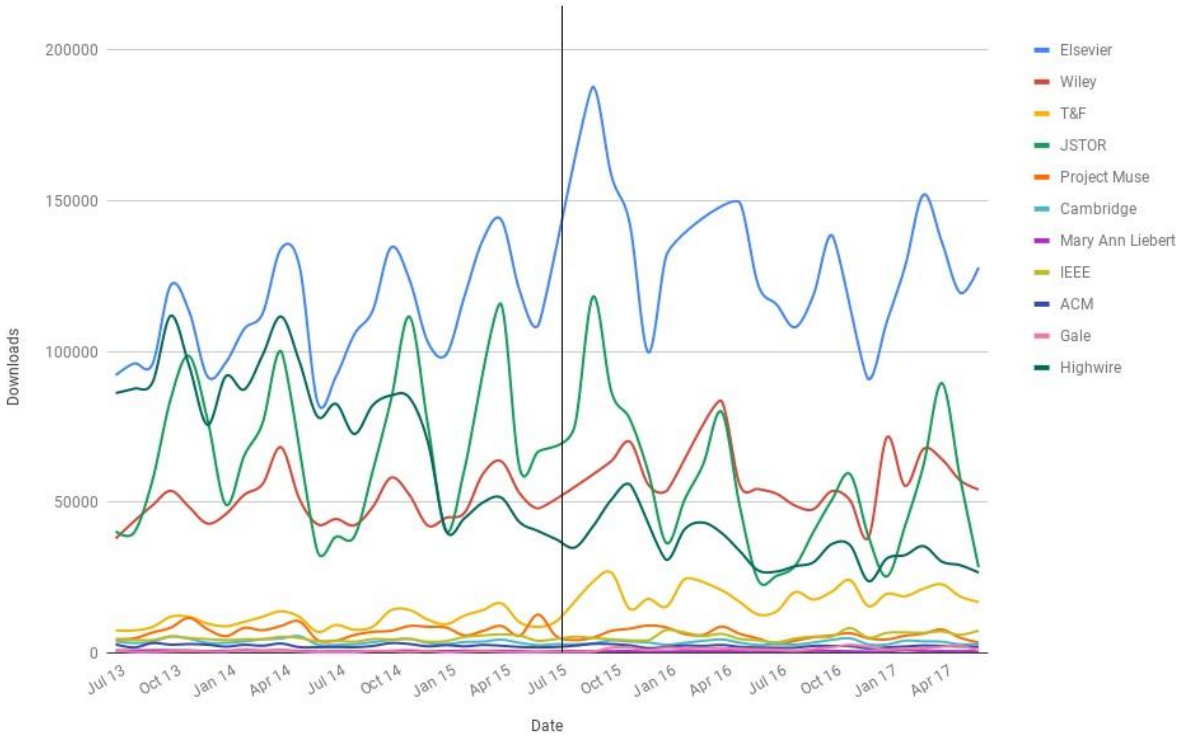

Figure 2. Ejournal platforms where usage changes had no correlation with EDS

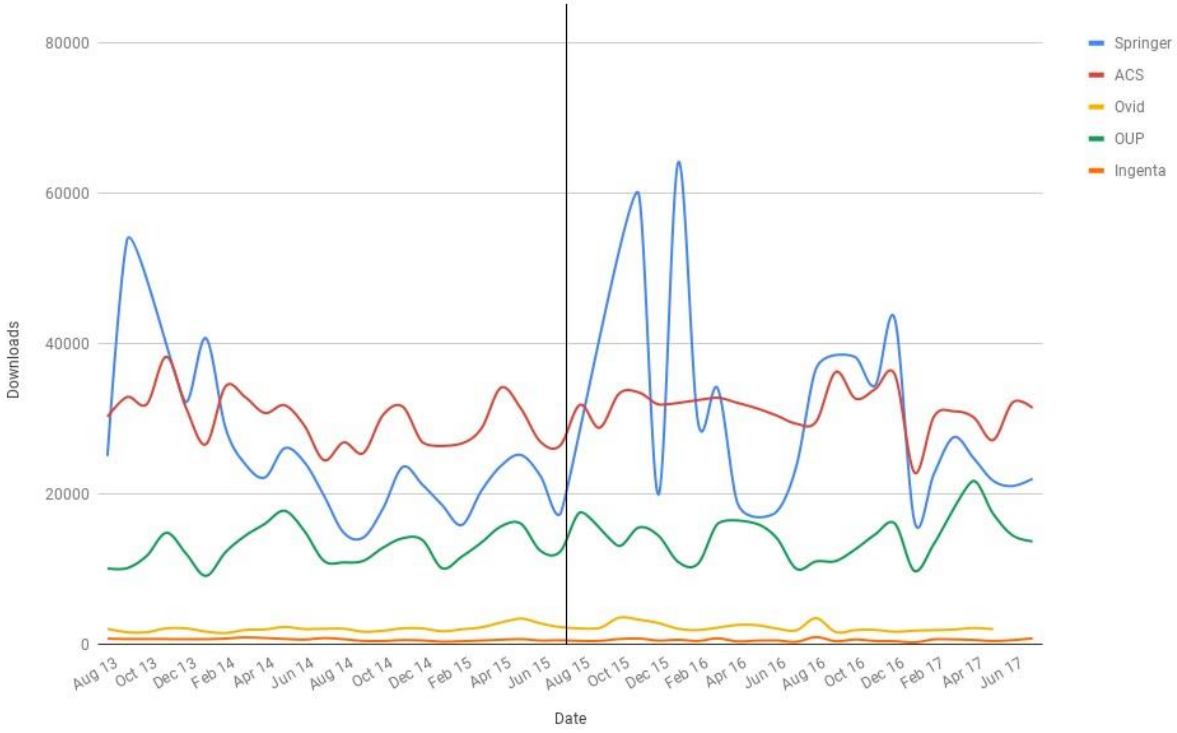

Figure 3. Usage for ebooks

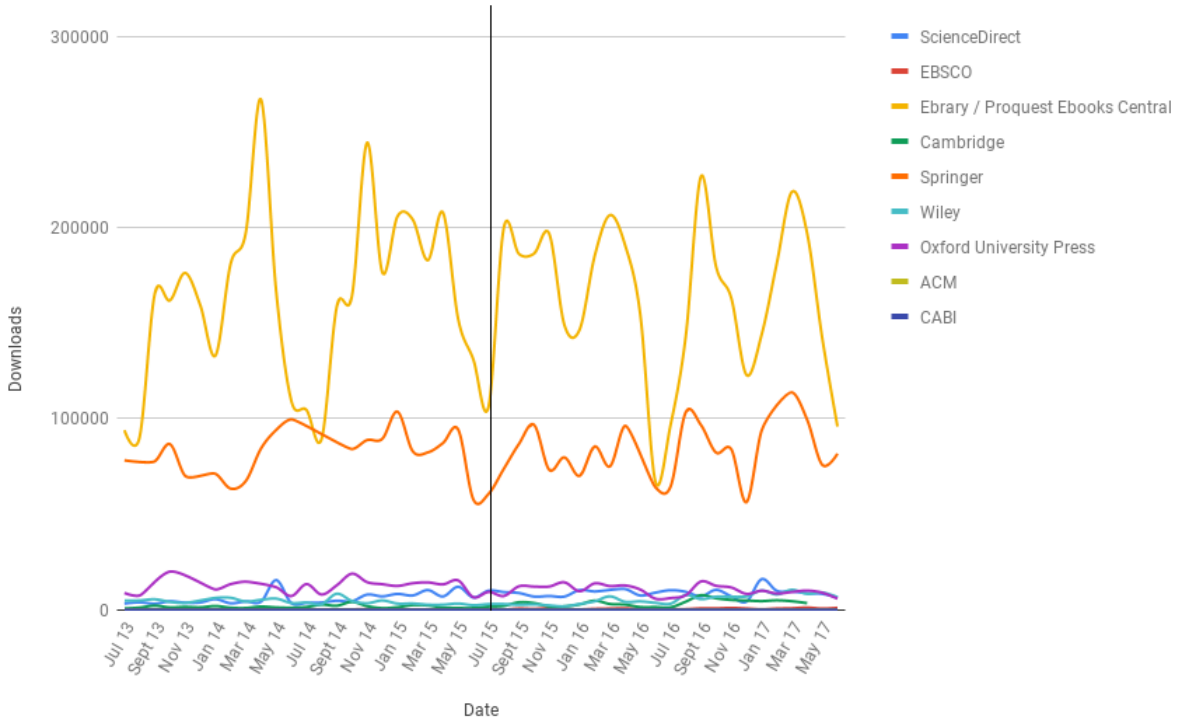

Figure 4. Usage for A&I databases